5

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/873,656	06/04/2001	Endong Xun	M61.12-0341	4270

7590 12/15/2003

Joseph R. Kelly
WESTMAN CHAMPLIN & KELLY
Suite 1600 - International Centre
900 South Second Avenue
Minneapolis, MN 55402-3319

EXAMINER	
PALMER, JUSTIN W	
ART UNIT	**PAPER NUMBER**
2654	3

DATE MAILED: 12/15/2003

Please find below and/or attached an Office communication concerning this application or proceeding.

PTO-90C (Rev. 10/03)

Office Action Summary	Application No.	Applicant(s)
	09/873,656	XUN ET AL.
	Examiner	Art Unit
	Justin Palmer	2654

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133).
- Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☐ Responsive to communication(s) filed on _____ .

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1-20* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1-20* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☒ The drawing(s) filed on *04 June 2001* is/are: a)☒ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

11)☐ The proposed drawing correction filed on _____ is: a)☐ approved b)☐ disapproved by the Examiner.

If approved, corrected drawings are required in reply to this Office action.

12)☐ The oath or declaration is objected to by the Examiner.

Priority under 35 U.S.C. §§ 119 and 120

13)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____ .

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

14)☐ Acknowledgment is made of a claim for domestic priority under 35 U.S.C. § 119(e) (to a provisional application).

a) ☐ The translation of the foreign language provisional application has been received.

15)☐ Acknowledgment is made of a claim for domestic priority under 35 U.S.C. §§ 120 and/or 121.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☒ Information Disclosure Statement(s) (PTO-1449) Paper No(s) 2 .
4) ☐ Interview Summary (PTO-413) Paper No(s). _____ .
5) ☐ Notice of Informal Patent Application (PTO-152)
6) ☐ Other: .

DETAILED ACTION

Claim Rejections - 35 USC § 103

1. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

2. This application currently names joint inventors. In considering patentability of the claims under 35 U.S.C. 103(a), the examiner presumes that the subject matter of the various claims was commonly owned at the time any inventions covered therein were made absent any evidence to the contrary. Applicant is advised of the obligation under 37 CFR 1.56 to point out the inventor and invention dates of each claim that was not commonly owned at the time a later invention was made in order for the examiner to consider the applicability of 35 U.S.C. 103(c) and potential 35 U.S.C. 102(e), (f) or (g) prior art under 35 U.S.C. 103(a).

3. Claims 1-20 are rejected under 35 U.S.C. 103(a) as being unpatentable over U.S. Patent 5,149,405 to Church in view of Manning and Schutze, *Foundations of Statistical Natural Language Processing*, MIT Press, 1999.

As per **claim 1**, Church discloses a method of processing a natural language input, comprising:

identifying a likely base noun phrase (baseNP) sequence based on POS sequences identified (see fig. 3, items 31 and 32; fig. 4) and;

outputting the likely baseNP sequence (see utilization step, fig. 2, item 25).

Church does not disclose identifying N-best part-of-speech (POS) sequences corresponding to the natural language input.

However, Manning et al. discloses identifying N-best part-of-speech (POS) sequences corresponding to the natural language input (see discussion of *k*-best tagging, p. 366).

At the time of invention it would have been obvious to one of ordinary skill in the art to modify Church to include using the N-best POS sequences as taught by Manning because tags that are likely in one situation could be equally likely in another situation, which is critical knowledge for constructing a parse, even shallow parses such as baseNP identification (see Manning, pp. 366-367).

As per **claim 2**, Church discloses a method of identifying a plurality of baseNP sequences for each of the N-best POS tag sequences (see algorithm describing probability search tree, col. 5, ll. 5-20); and

calculating which of the plurality of baseNP sequences is most likely (ibid, wherein the probability search tree contains rankings of the most likely baseNP sequences).

As per **claim 3**, Church discloses a method of calculating which of the plurality of baseNP sequences is most likely comprising:

calculating a likely baseNP sequence that is most likely based on lexical information indicative of a position of words in the natural language input relative to baseNPs identified in the baseNP sequences (see discussion of precedence of parsing statistics, col. 9, ll. 1-55, where the bracketings denote word positions and baseNP sequences already identified).

As per **claim 4**, Church discloses a method of calculating a likely baseNP sequence that is most likely based on lexical information comprising:

calculating a likely baseNP sequence that is most likely based on lexical information indicative of POS tags assigned to the words in the natural language input (see col. 9, ll. 9-14).

As per **claim 5**, Church discloses a method of calculating a likely baseNP sequence based on lexical information.

Church does not disclose calculating a likely baseNP sequence that comprises:

calculating a likely baseNP sequence based on the lexical information for every word in the natural language input.

However, Manning et al. discloses using lexical information for every word in the natural language input in part-of-speech tagging (see discussion of lexical information, p. 344).

At the time of the invention it would have been obvious for one of ordinary skill in the art to modify Church to take into account lexical information for every word in the natural language input when identifying baseNP sequences because it helps pinpoint a word's part-of-speech when the distribution of a word's usages across different parts of speech is uneven, a typical occurrence (see Manning p. 344).

As per **claim 6**, Church discloses the method of calculating which of the plurality of baseNP sequences is most likely comprises:

calculating which of the plurality of baseNP sequences is most likely over the entire sentence (see col. 9, l. 60-col. 10, l. 7).

As per **claim 7**, Church discloses a method wherein a baseNP rule comprises a sequence of POS tags corresponding to words in the natural language input identified as a baseNP and wherein calculating a likely baseNP sequence comprises:

calculating a probability of POS tags and baseNP rules, given their context (see col. 3, ll 49-58 and subsequent detailed description).

As per **claim 8**, Church discloses a method wherein calculating a probability of POS tags and baseNP rules comprises:

calculating the probability of POS tags and baseNP rules given n prior POS tags or baseNP rules (see trigram discussion, col. 4, ll. 16-23).

As per **claim 9**, Manning et al. discloses a method wherein identifying a likely baseNP sequence includes:

calculating a probability of each of the N-best POS sequences given the natural language input (see discussion of *k*-best, p. 366).

As per **claim 10**, Church discloses a natural language processing system for processing a natural language input, comprising:

a base noun phrase (baseNP) identifier configured to receive part-of-speech (POS) tag sequences for the natural language input and identify a likely baseNP sequence of baseNPs corresponding to the natural language input, given the POS tag sequences.

Church does not disclose identifying N-best part-of-speech (POS) sequences corresponding to the natural language input.

However, Manning et al. discloses identifying N-best part-of-speech (POS) sequences corresponding to the natural language input (see discussion of *k*-best tagging, p. 366).

At the time of invention it would have been obvious to one of ordinary skill in the art to modify the system disclosed by Church to include using the N-best POS sequences as taught by Manning because tags that are likely in one situation could be equally likely in another situation, which is critical knowledge for constructing a parse, even shallow parses such as baseNP identification (see Manning, pp. 366-367).

As per **claim 11**, Church does not disclose a POS tagger, coupled to the baseNP identifier, receiving the natural language input and calculating the N-best POS tag sequences corresponding to the natural language input.

However, Manning et al. discloses identifying N-best part-of-speech (POS) sequences corresponding to the natural language input (see discussion of *k*-best tagging, p. 366).

At the time of invention it would have been obvious to one of ordinary skill in the art to modify the system disclosed by Church to include using the N-best POS sequences as taught by Manning because tags that are likely in one situation could be equally likely in another situation, which is critical knowledge for constructing a parse, even shallow parses such as baseNP identification (see Manning, pp. 366-367).

As per **claim 12**, Church discloses a system wherein the baseNP identifier is configured to identify a plurality of baseNP sequences for each of the N-best POS tag sequences (see algorithm describing probability search tree, col. 5, ll. 5-20) and calculate which of the plurality of baseNP sequences is most likely (ibid, wherein the probability search tree contains rankings of the most likely baseNP sequences).

As per **claim 13**, Church discloses a system wherein the baseNP identifier further comprises:

a unified statistical model that includes lexical information indicative of a position of words in the natural language input relative to baseNPs identified in the baseNP sequences (see discussion of precedence of parsing statistics, col. 9, ll. 1-55, where the bracketings denote word positions and baseNP sequences already identified).

As per **claim 14**, Church discloses a system wherein a baseNP rule comprises a sequence of POS tags corresponding to words in the natural language input that are identified as a baseNP and wherein the unified statistical model includes a baseNP rule component for calculating a probability of POS tags and baseNP rules, given contextual information (see col. 3, ll 49-58 and subsequent detailed description).

As per **claim 15**, Church discloses a system wherein the baseNP rule component is configured to calculate the probability of POS tags and baseNP rules, given n prior POS tags and baseNP rules.

As per **claim 16**, Church discloses a system wherein the natural language input comprises a sentence and wherein the unified statistical model is configured for calculating which of the plurality of baseNP sequences is most likely over the entire sentence (see col. 9, l. 60-col. 10, l. 7).

As per **claim 17**, Church discloses a method of processing a linguistic input, comprising:

identifying one or more base noun phrases (baseNPs) for each of the POS sequences to form a plurality of different possible baseNP sequences corresponding to each of the POS sequences identified (see fig. 3, items 31 and 32; fig. 4);

for each baseNP sequence, identifying whether it is a likely baseNP sequence based on a probability of the associated POS sequence and a probability of the baseNP sequence, given lexical information indicative of a position of words in the linguistic input relative to the baseNPs identified in the baseNP sequence (see col. 9, ll. 20-35 where noun phrase probabilities are noted; col. 7, ll. 1-49 where POS sequence probabilities are noted); and

outputting the likely baseNP sequence identified (see utilization step, fig. 2, item 25).

Church does not disclose identifying N-best part-of-speech (POS) sequences corresponding to the natural language input.

However, Manning et al. discloses identifying N-best part-of-speech (POS) sequences corresponding to the natural language input (see discussion of *k*-best tagging, p. 366).

At the time of invention it would have been obvious to one of ordinary skill in the art to modify Church to include using the N-best POS sequences as taught by Manning because tags that are likely in one situation could be equally likely in another situation, which is critical knowledge for constructing a parse, even shallow parses such as baseNP identification (see Manning, pp. 366-367).

As per **claim 18**, Church discloses a method wherein identifying one or more baseNPs for each of the N-best POS sequences comprises:

identifying baseNP rules for a POS sequence, the baseNP rules comprising a sequence of POS tags corresponding to words in the linguistic input, identified as a baseNP (see stochastic parser, which contains rules for identifying NP from a POS sequence, col. 9, ll. 36-38).

As per **claim 19**, Church discloses the method wherein identifying whether each baseNP sequence is a likely baseNP sequence, comprises:

calculating a probability of POS tags and baseNP rules, given n prior POS tags or baseNP rules in the POS sequence (see trigram, col. 4, ll. 16-20).

As per **claim 20**, Church discloses a method wherein the linguistic input comprises a sentence and wherein identifying whether each baseNP sequence is a likely baseNP sequence comprises:

identifying whether each baseNP sequence is a likely baseNP sequence over the entire sentence (see col. 9, l. 60-col. 10, l. 7).

Conclusion

4. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure. U.S. Patent Application Publication 2002/0077806 to Tarbouriech et al.; U.S. Patent 6,278,967 B1 to Akers et al.; U.S. Patent 6,631,346 to Karaorman et al.; U.S. Patent 5,696,962 to Kupiec; U.S. Patent 5,890,103 to Carus; U.S. Patent 6,167,368 to Wacholder; U.S. Patent 5,799,269 to Schabes et al.; U.S. Patent 5,930,746 to Ting; U.S. Patent 5,963,940 to Liddy et al.; U.S. Patent 6,182,028 B1 to Karaali et al.; U.S. Patent 6,289,304 B1 to Greffenstette; Brill, Eric. "Unsupervised learning of disambiguation rules for part of speech tagging." In *WVLC 3*, p. 1-13, 1995; Brill, Eric. "Transformation-based error-driven parsing." In *Proceedings Third International Workshop on Parsing Technologies*, Tilburg/Durbuy, The Netherlands/Belgium, 1995; Abney, Steven. "Part-of-speech tagging and partial parsing." In Steve Young and Gerrit Bloothooft (eds.), *Corpus-Based Methods in Language and Speech Processing*, pp. 118-136. Dordrecht: Kluwer Academic, 1996; all disclose methods of combing the using POS tagging in NP identification systems, co-training POS taggers with other machine learning tasks, and the similarities in the problems of partial (or shallow) parsing and POS tagging.

5. Any inquiry concerning this communication or earlier communications from the examiner should be directed to Justin Palmer whose telephone number is (703) 305-8663. The examiner can normally be reached on Monday-Thursday 7:00 AM-5:15 PM Eastern.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Talivaldis Ivars Smits can be reached on (703) 306-3011. The fax phone number for the organization where this application or proceeding is assigned is (703) 872-9306.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Talivaldis Ivars Smits can be reached on (703) 306-3011. The fax phone number for the organization where this application or proceeding is assigned is (703) 872-9306.

Any inquiry of a general nature or relating to the status of this application or proceeding should be directed to the receptionist whose telephone number is (703) 305-4700.

Jwp.
December 8, 2003

TĀLIVALDIS IVARS ŠMITS
PRIMARY EXAMINER